Defining Acceptable Thresholds for PROMs

in Total Hip and Knee Replacement

J. Christiaan Keurentjes, MD PhD

16th of November 2014

1 Clinical Background

2 Focus on Improvement

3 Key Assumptions

Joint Degeneration

- 90% primary osteoarthritis
- 8 million patients in UK
- Hands, knees, hips
- Symptoms: pain, stiffness, joint effusion
- Natural history: slowly progressive

Treatment options

1. Exercise
2. Weight loss
3. Painkillers
4. Corticosteroid injections
5. Surgery

Principles of Surgery

- Remove damaged joint
- Replace with either:
  - Artificial Joint $\rightarrow$ Joint Replacement
  - Bone $\rightarrow$ Arthrodesis

Joint Replacement: Complications

- Early: *up to 2 years*
  - Medical: MI, CVA, VTE
  - Infection: 1–2%
  - Technical / Mechanical: $<1\%$
- Long term: *longer than 10 years*
  - Loosening: $\leq 5\%$ at 10 years follow-up

Joint Replacement: Outcome

- Alleviate pain
- Improve Functioning
- Improve Health-Related Quality-of-Life (HRQoL)

Joint Replacement: Outcome

- On average large improvement in HRQoL
- But not for all patients:
  - 0–20% persisting pain
  - 10–30% unsatisfied with results
- Limited therapeutic options in failed cases

1 Clinical Background

2 Focus on Improvement

3 Key Assumptions

Clinically Important Differences

- Thresholds indicating *relevant* improvement
- Dichotomous outcome
  - Responder:

    $$PROM_{outcome} - PROM_{baseline} > CID$$

  - Non-responder:

    $$PROM_{outcome} - PROM_{baseline} \leq CID$$

CID: Anchor-based

- Ingredients:
  - $PROM_{outcome}$, $PROM_{baseline}$
  - Relevant anchor-question at outcome, typically Likert-scale
- CID: Mean improvement of patients, who experienced relevant improvement according to Anchor-question

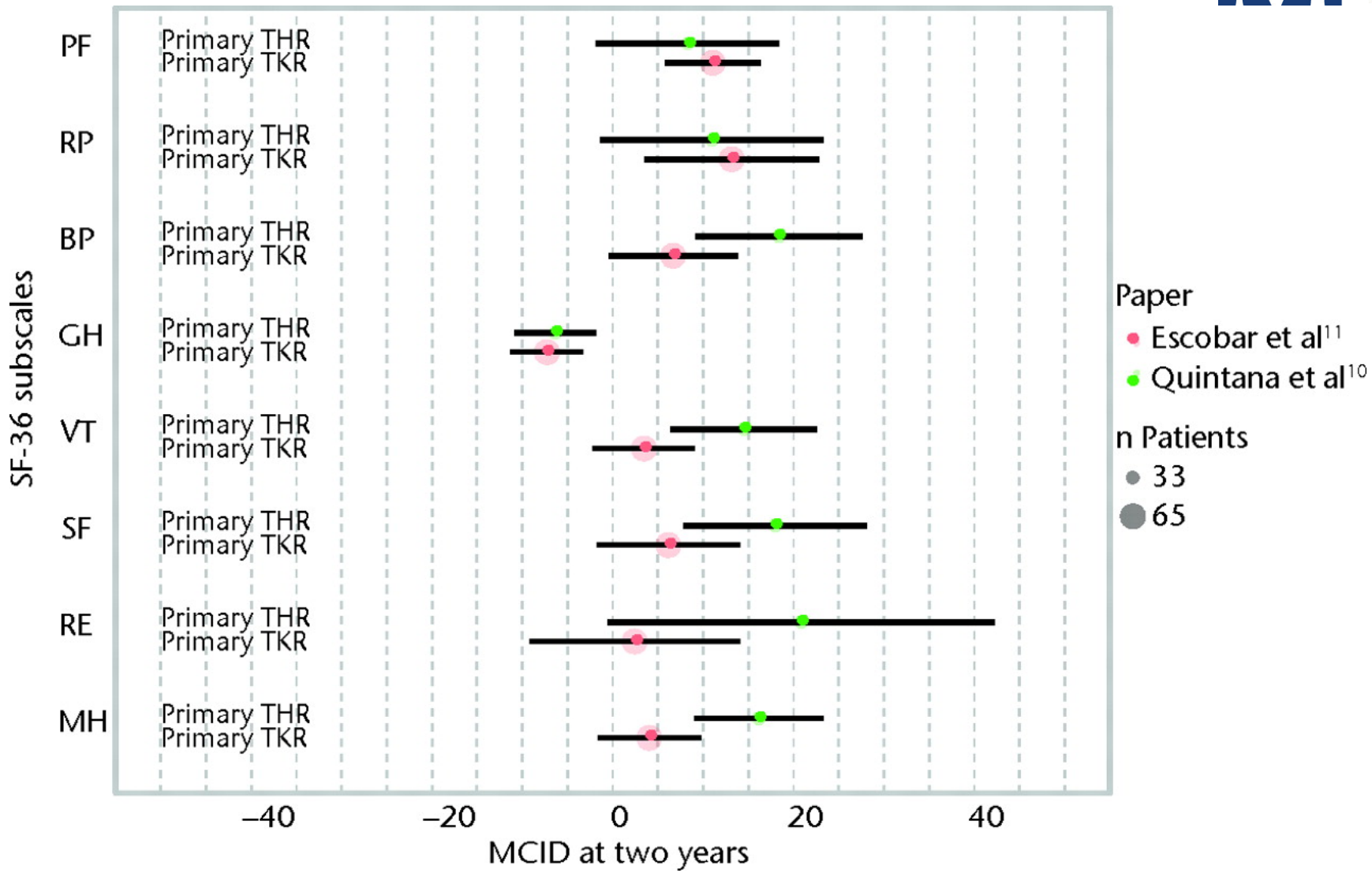
CID: Anchor-based: Quintana et al

- Ingredients:
  - $SF36_{outcome}$ at 2 years follow-up, $SF36_{baseline}$
  - 5-point Likert-scale:
    1. "a great deal better"
    2. **"somewhat better"**
    3. "equal"
    4. "somewhat worse"
    5. "a great deal worse"

CID: Anchor-based: Quintana et al

- 469 patients at baseline, 310 at follow-up

	n	%
"a great deal better"	254	81.9
"somewhat better"	**33**	**10.6**
"equal"	8	2.6
"worse"	10	3.2


SF-36 subscales
PF
RP
BP
GH
VT
SF
RE
MH
Primary THR
Primary TKR
−40
−20
0
20
40
MCID at two years
Paper
Escobar et al[11]
Quintana et al[10]
n Patients
33
65

CID: Anchor-based

- Clinically meaningful
- Arbitrariness in anchor
- Impractical approach in treatments with large effect size

CID: Distribution-based

- Ingredients:
  - $SD(PROM_{baseline})$
  - "Standard" effect size estimate (ES):

    e.g. Cohen's 0.2, 0.5 or 0.8
- CID: $SD(PROM_{baseline}) * ES$

CID: Distribution-based

- No apparent clinical meaning of estimate
- Arbitrariness in effect size
- Makes use of data of all patients

CID: Our Approach

1. Distribution-based estimation
2. Anchor-based validation

CID: Our Approach

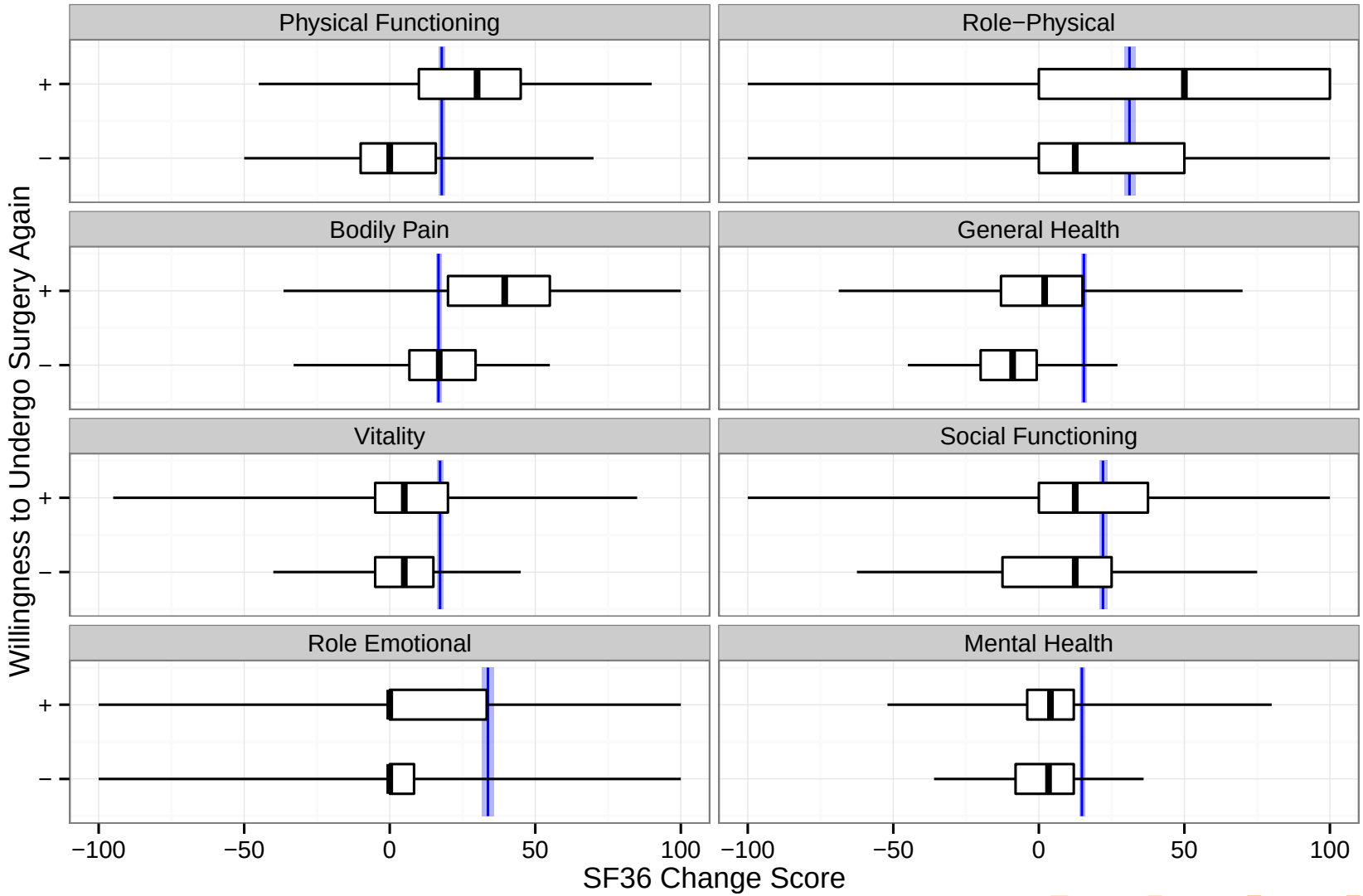
1. Distribution-based estimation:

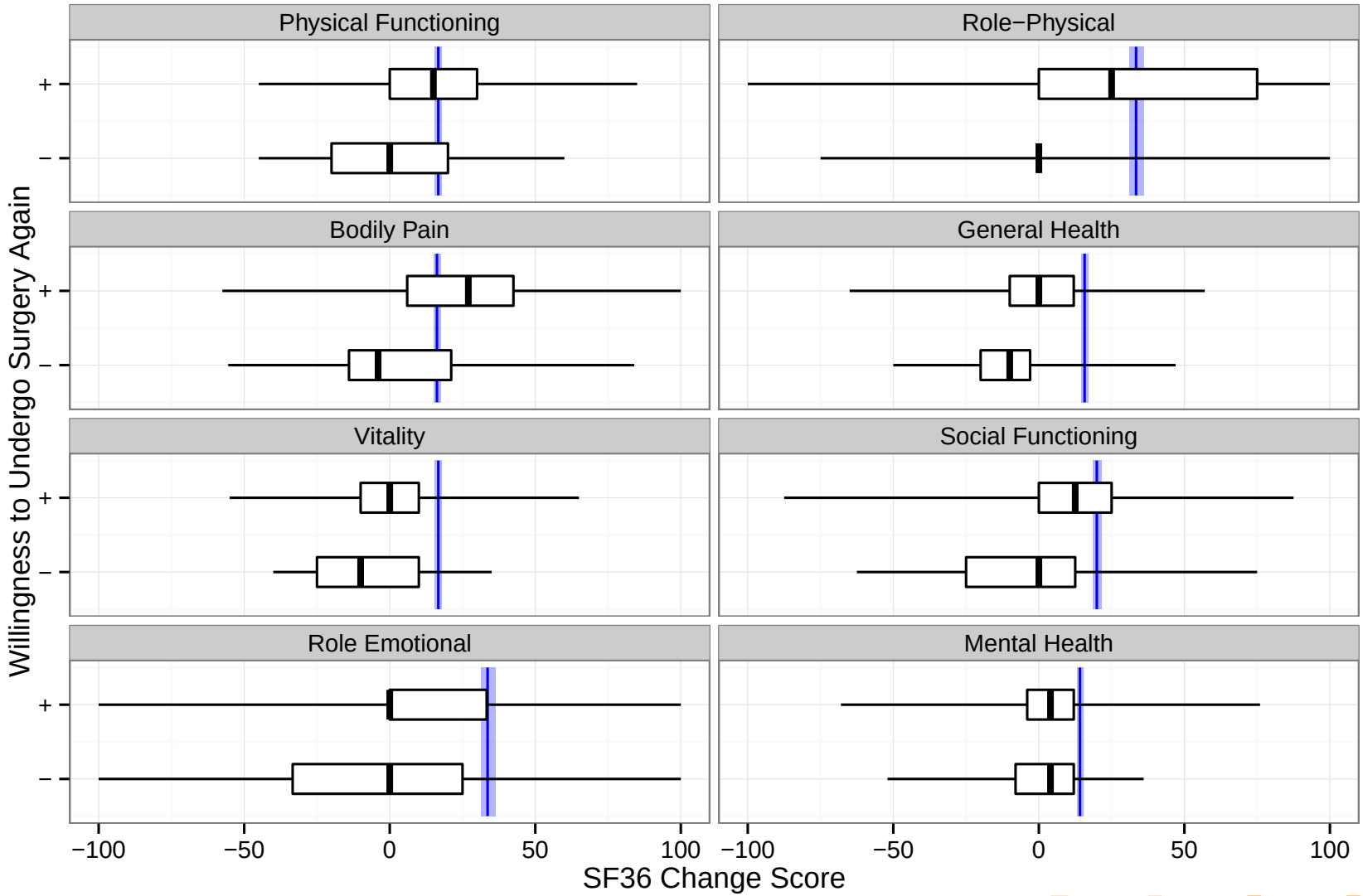
   $SD(SF36_{baseline}) * 0.8$

2. Anchor-based validation:

   Cross-tabulation with relevant overall dichotomous outcome:

   "Willingness to undergo surgery again"


Physical Functioning
Role−Physical
Bodily Pain
General Health
Vitality
Social Functioning
Role Emotional
Mental Health
Willingness to Undergo Surgery Again
+
−
−100
−50
0
50
100
SF36 Change Score


Physical Functioning
Role−Physical
Bodily Pain
General Health
Vitality
Social Functioning
Role Emotional
Mental Health
Willingness to Undergo Surgery Again
+
−
−100
−50
0
50
100
SF36 Change Score

CID: Odds Ratio

SF36 Subscale	Hip Replacement	Knee Replacement
Physical Functioning	5.86	1.80
Role Physical	2.08	2.98
Bodily Pain	3.30	4.72
General Health	4.92	1.26

CID: Odds Ratio

SF36 Subscale	Hip Replacement	Knee Replacement
Vitality	1.11	0.78
Social Functioning	1.89	3.35
Role Emotional	2.84	0.68
Mental Health	1.06	0.95

1 Clinical Background

2 Focus on Improvement

3 Key Assumptions

1st CID Assumption

- Patients with a high score at baseline cannot attain a relevant improvement:

$$PROM_{max} - PROM_{baseline} \leq CID$$

1st CID Assumption

- SF36 Bodily Pain
- Range: 0 – 100
- 0: worst possible pain, 100: no pain at all
- CID: 16.8

1st CID Assumption

- Patient A: no "pain", only stiffness and functional limitation
- $BP_{baseline} = 100$
- $p(BP_{outcome} - BP_{baseline} > 16.8) = 0$
- **Makes sense**

2nd CID Assumption

- CID is equal for all
- Questionable due to regression to the mean: patients with worse preoperative PROM are more likely to improve more than patients with less extreme preoperative PROM

2nd CID Assumption

- CID is equal for all
- Possible solutions:
  - Separate CIDs for strata of $PROM_{baseline}$?
  - Adjust for $PROM_{baseline}$ in multivariate model?

Paper

JC Keurentjes, M Fiocco, RG Nelissen.

Willingness to undergo surgery again validated clinically important differences in health-related quality of life after total hip replacement or total knee replacement surgery.

Journal of Clinical Epidemiology, 2014: p114-120.